AD-A106 138   NAVAL POSTGRADUATE SCHOOL MONTEREY CA   F/G 5/8
                THE EFFECTS OF CERTAIN BACKGROUND NOISES ON THE PERFORMANCE OF --ETC(U)
                SEP 80   R S ELSTER
UNCLASSIFIED   NPS54-80-010   NL

| OF |
AF
A106138

END
DATE
FILMED
11 R
DTIC

LEVEL II ②

NPS54-80-010

AD A106138

DTIC FILE COPY

NAVAL POSTGRADUATE SCHOOL
Monterey, California

DTIC
S ELECT
OCT 27 1981
E

The Effects of Certain Background Noises
on the Performance of a Voice
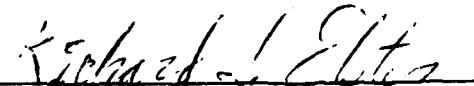Recognition System

by

Richard S. Elster

September 1980

81 10 26 034

NAVAL POSTGRADUATE SCHOOL

Monterey, California

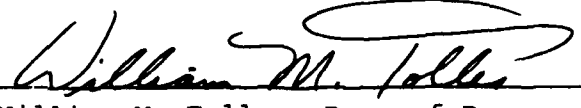Rear Admiral J. J. Ekelund
Superintendent

D. A. Schrady
Acting Provost

Richard S. Elster, Professor

Reviewed by: Released by:

Carl R. Jones, Chairman   William M. Tolles, Dean of Research
Dept. of Administrative
Sciences

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
NPS54-80-010	AD-A106138	

4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED
The Effects of Certain Background Noises on the Performance of a Voice Recognition System.	Technical
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)
Richard S. Elster	

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Naval Postgraduate School Monterey, CA 93940	N0003980WR09047

11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE
Naval Postgraduate School Monterey, CA 93940	September 1980
	13. NUMBER OF PAGES

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report)
Naval Electronics Systems Command Washington, D.C. 20360	Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, If different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

VTAG, Voice Recognition, Automatic Word Recognition.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report describes an experiment concerning the influence of different levels of background noise on the performance of a voice recognition system. Three levels of noise were used: 38, 65 and 75dBA.

The results suggest that if the device is trained at 38, 65, or 75dBA, performance will be satisfactory in 38 or 65dBA environments, while to be used in a 75dBA environment, the device should

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73
(Page 1) S/N 0102-014-6601 |

be trained in a 65 or 75dBA environment,

DD Form 1473
1 Jan 73
S N 0102-014-6601
(Page 2)

2

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

TABLE OF CONTENTS

LIST OF FIGURES

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

11

LIST OF TABLES

FOREWORD

This investigation was sponsored by Mr. C. C. Stout, NAVELEX, Code 330. The work was performed by the investigator at the Naval Postgraduate School, Monterey, California.

This report is the second in a series concerned with the possible applications of voice recognition technology in command and control tasks. The first report was, "Experiments with Voice Input for Command and Control: Using Voice Input to Operate a Distributed Computer Network," (Technical Report NPS55-80-016), by Gary K. Poock, April 1980.

THE EFFECTS OF CERTAIN BACKGROUND NOISES

ON THE PERFORMANCE OF A VOICE RECOGNITION SYSTEM

I.   EXECUTIVE SUMMARY

In this experiment the performance of a voice recognition device was examined as a function of background noise conditions. A subject trained the recognizer in one background noise condition and used it in three background noise conditions.

The most important findings were that if the voice recognition device is to be used in a 75dBA conversational noise environment, then training the system in a 65 or 75dBA conversational environment will yield fewer errors than when it is trained in a 38dBA white noise environment; while if one trains in a 38, 65, or 75dBA, performance will be satisfactory when used in 38 or 65dBA environments.

II. INTRODUCTION

A.   Problem Voice recognition equipment is being considered for use in various military command and control functions. The effects, if any, of background noises upon the performance of a command and control system using voice recognition equipment are largely unknown. Before voice recognition equipment is used in operational command and control systems, the relationships between system performance and background noise must be understood.

B.   Objective The objective of the experiment described in this report was to determine the effect of background noise,

4

including human conversation, on the performance of a voice recognition system.

C.   <u>Background</u>   Technology allowing the use of voice input to control machines has recently been developed.   Although in relative infancy, this technology has yielded equipment that can be trained to recognize a set of utterances from nearly continuous speech.   Applications and experiments using voice recognition equipment are burgeoning.   Poock (1980), for instance, reported on the use of voice input to operate a distributed computer network.   Also in 1980, the Department of Defense (DoD) sponsored a conference on voice interactive systems (Voice Interactive Systems: Applications and Payoffs, 1980).   The DoD conference featured three days of presentations covering a number of ways in which voice technology can be used in man-machine systems.   A presentation at the DoD conference by Thomas G. Drennen discussed the effect of attack/ fighter cockpit noise on speech characteristics and on voice recognition system performance.   Drennen reported that the voice recognition system he used performed more accurately under extremely high (106 or 114dB) noise levels when the training had been under similar noise levels (114dB) than when the training had been done at low (10dB) noise level.[1] At testing levels of 10 or 101dB, however, recognition accuracy was higher if the training had been done in a 10 rather than a 114dB environment.

5

Drennen's noise environment represented cockpit conditions under different aircraft power settings. In many command and control applications, background voice messages and conversations are present and might influence the performance of voice recognition devices. It is important to determine if Drennen's findings extend to environments in which the background noise is human speech and to less extreme dB levels of background noise.

III. APPROACH

A.   Experimental Setting The experiment was conducted in a soundproof chamber. A model T600 Threshold Technology, Inc. voice recognition device was used with a Shure model SM10 microphone. With added memory modules, up to 256 two-second voice utterances could have been used. In this experiment, 50 utterances were used. A maximum utterance length of two seconds was a limitation imposed by the voice recognition device. For more details on the operation of voice recognition equipment, see Poock (1980).

B.   Independent Variables Two independent variables were investigated in this experiment: first, the level of background noise during the training of the model T600 voice recognition device; second, the level of background noise during the testing of the voice recognition device. The training noise level and the testing noise level independent variables had the same three levels of noise: ambient noise (an average of about 48dBA), conversational noise at an average of 65dBA,

6

and conversational noise at an average of 75dBA.[2] For both the 65dBA and 75dBA average noise levels, the sound levels varied from the average value by no more than ±7dBA. Sound levels were measured at the microphone connected to the voice recognition device.

The levels of background noise were measured using the dBA-weighting network. The A-weighting network is very good at giving a quick estimate of the interference of noise upon speech (MIL-HDBK-759, p. 358). When dBA levels of 90-95dBA and greater were tried, the voice recognizer tended to emit a nearly continuous string of extra outputs even though no one was speaking to it. Therefore, background noises of that level were not considered for use in this experiment. Speech interference levels (SIL) are often used to estimate maximum permissable levels of background noises (Bragdon, p. 79). The SIL can be determined from the dBA-weighted network (Bragdon, p. 79). Tables are available (see, for instance, the Human Engineering Guide to Equipment Design, p. 193) demonstrating the relationship between speech level (normal, raised, very loud, and shouting), distance between talker and listener, and level of background noise that barely permits reliable conversation. For example, for reliable conversation when the speaker is one foot from the listener, the background noise should not exceed 75dBA. A background noise of 65dBA or less should permit reliable conversation when the speaker and listener are three feet apart. Bragdon (1971, p. 79)

7

reports that when background noise approaches 80dBA, hearing accuracy declines. Bragdon (1971, p. 80) also describes a survey which found that 71dBA was a maximum acceptable level for background noise for voice communications. At noise levels greater than that, people reported their job performance was adversely impacted. In conclusion, the three levels of background sound used in this experiment (38dBA, 65dBA, and 75dBA) should have covered the range of background noise intensities likely to be found in many command and control environments.

C. Dependent Variables. Three types of voice recognition system errors were recorded and added together to form the error measure used in the analysis of results of the experiment.

- Wrong outputs: the recognizer gave the wrong response to the subject's utterance.

- "Beeps": the Model T600 Threshold Technology, Inc. voice recognition device emitted an audible beep when it did not recognize an utterance.

- Extra outputs: the voice recognition device emitted a response when the subject had not emitted an utterance. These outputs could occur when the microphone was open either before or after an utterance.

The dependent variable used in the analysis was formed by summing together the number of errors made by the voice recognition device in each subject x test condition combination.

D. Experimental Design. This was a two-factor experiment with repeated measures on one factor (Winer, p. 302); Subjects were nested within one factor. Each subject trained the voice

recognition device under one of the noise conditions and tested
the voice recognition device under each of the three noise con-
ditions. Six subjects were randomly assigned to each of the
three training conditions. The ordering of presentation of
the test conditions was done such that each test condition
appeared an equal number of times in first, second, and third
place for each training condition. Figure 1 portrays the de-
sign of the experiment.

E.   Training and Testing. Each subject trained the voice
recognition device to the same list of 50 utterances. (A copy
of the list of utterances is provided in Appendix I).

During the training phase, the subject would repeat each
utterance 10 times. Following the 10 repetitions of an utter-
ance, the device was deemed to be trained if the utterance
was recognized correctly two out of three times. Training
with an utterance continued until the two-out-of three cri-
terion was satisfied.

During the testing phase of the experiment, the subject
was instructed to read each word only once (under each test
background noise condition). An error was counted if the voice
recognizer emitted the wrong output, "beeped", or emitted an
output when the subject had not spoken one of the utterances.

A copy of the instructions given to the subjects is
given in Appendix II. The instruction sheet also includes
prompts to be followed by the experimenter.

9

Figure 1

CONCEPTUAL DESIGN OF THE EXPERIMENT

Training Noise Level (dBA)

	38	65	75
	Subject # 1 ●●●●6	Subject # 7 ●●●●12	Subject # 13 ●●●●18
Testing Noise Level (dBA) 38	×		
65	×		
75	×		

Each subject trained the device in one dBA level, but tested device in all three dBA levels.

EXAMPLE:
Recognition errors with subJ #18 when device was trained at 75 dBA and tested at 75 dBA

10

IV RESULTS

A. Number of Errors. Table 1 presents the number of errors, and mean number of errors for the different experimental conditions. (Appendix III presents the data by type and number of errors, by subject.)

B. Analysis of Variance. An analysis of variance was made of the error data shown in Table 1. Table 2 presents the results of that analysis.

The only F-statistic significant in Table 2 is the one for test noise level. Because certain assumptions about the subjects' covariance matrices must be met or the sampling distribution of the F statistic will not be the F distribution, a conservative test was also applied to the Test Noise Level variable. Winer (pp. 305-306) describes a conservative test developed by Greenhouse and Geisser. For that test, the degrees of freedom to be used in this experiment for the critical value of the F statistic for the Test Noise Level are $(1,15)$. Using those degrees of freedom, the F statistic for Test Noise Level is still statistically significant ($p < .01.$).

Scheffé's confidence intervals (Winer, p. 85) were used to make a posteriori comparisons among the three testing noise condition means. The confidence intervals are presented in Table 3.

The results in Table 3 indicate (because zero is outside the intervals) that the number of errors made by the voice

11

TABLE 1

Errors by Training Level, Testing Level, and by Subject

		Noise Level in Training																	
		38dBA						65dBA						75dBA						Overall Mean for a Testing Noise Level
SUBJECT		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
38dBA		0	0	2	4	0	1	2	0	2	4	2	1	1	5	1	4	0	2	$\bar{X} = 1.72$
		$\bar{X}_{11} = 1.16$						$\bar{X}_{12} = 1.83$						$\bar{X}_{13} = 2.16$						
65dBA		2	1	2	0	1	0	1	0	1	2	3	0	0	3	1	0	0	1	$\bar{X} = 1.00$
		$\bar{X}_{21} = 1.0$						$\bar{X}_{22} = 1.16$						$\bar{X}_{23} = .83$						
75dBA		6	25	4	15	7	6	8	3	3	7	16	8	1	3	4	16	3	6	$\bar{X} = 7.83$
		$\bar{X}_{31} = 10.5$						$\bar{X}_{32} = 7.5$						$\bar{X}_{33} = 5.5$						
Overall Mean For a Training Noise Level		$\bar{X} = 4.22$						$\bar{X} = 3.50$						$\bar{X} = 2.83$						$\bar{X} = 3.52$ Grand Mean

12

TABLE 2

Summary of the Analysis of Variance of the Voice Recognition Errors

Source of Variance	df	MS	F
Between Subjects	17		
Training Noise Level	2	8.68	.62
Subjects within Training Level	15	13.87	—
Within Subjects	36		
Test Noise Level	2	253.68	17.51*
Test Level x Training Noise Level	4	15.52	1.07
Test Level x Subjects within training level	30	14.48	—

* $p \ll .01$.

13

TABLE 3

Scheffé's Confidence Intervals for Differences Between
Pairs of Testing Condition Means

Contrast	Difference Between the Sample Means	95% Confidence Interval
$\bar{\mu}_{38} - \bar{\mu}_{65}$	$1.72 - 1.00 = .72$	$C \ [-2.54 \leq \bar{\mu}_{38} - \bar{\mu}_{65} \leq 3.98] = .95$
$\bar{\mu}_{75} - \bar{\mu}_{38}$	$7.83 - 1.72 = 6.11$	$C \ [2.85 \leq \bar{\mu}_{75} - \bar{\mu}_{38} \leq 9.37] = .95$
$\bar{\mu}_{75} - \bar{\mu}_{65}$	$7.83 - 1.00 = 6.83$	$C \ [3.57 \leq \bar{\mu}_{75} - \bar{\mu}_{65} \leq 10.09] = .95$

recognition device under the 75dBA testing noise level was significantly greater than the number of errors made under either the 65dBA or the 38dBA testing levels. The confidence interval in Table 3 for the 38dBA vs. 65dBA contrast shows (because zero is inside the interval) that the numbers of errors made under those two testing conditions were not significantly different.

Figure 2 provides plots of the average number of errors made by the recognition device for each test noise level at each of the training noise levels. In Figure 2, the noise levels are reported in terms of decibels, while in Figure 3 the sound pressure levels are presented in terms of microbars. The average threshold for human hearing is .0002 microbars which equals .0002 dynes/cm^2 or 10^{-6} watts/cm^2 (Woodworth and Schlosberg, p. 325). The microbar levels were found by solving equation 1 for P when SPL (sound pressure level) was 38, 65, or 75dB, and P_0 = .0002.

Equation 1 $SPL = 20 \log_{10} \dfrac{P}{P_0}$

The lines graphed on Figure 2 and 3 have rather different appearances because of the logarithmic relationship between the decibel scale and sound pressures.

Statistically significant ($p \leq .05$) differences between pairs of training x testing condition mean numbers of errors are indicated on Figure 2. Scheffé's confidence intervals were used with α = .05 to contrast pairs of test x training condition means.[3,4,5]

15

Figure 2.

MEAN NUMBER OF RECOGNITION DEVICE ERRORS BY TEST AND TRAINING NOISE LEVELS IN dBA.

Significant <u>a posteriori</u>
Contrasts ($p \leq .05$):

a_1 vs. a_2
a_1 vs. a_2
a_1 vs. a_3
a_1 vs. a_3 (Average of tested at 38 or 65; trained at 38)

10.5
a_1

Tested at 75 dBA

7.5

5.5

O_1 — Overall 75 dBA Test Condition Mean

Overall 38 dBA Test Condition Mean
O_2

Overall 65 dBA Test Condition Mean
O_3

2.16

1.16
a_3
Tested at 38 dBA
1.83

Tested at 65 dBA
.83

a_2 —1.0

1.16

Mean Number of Errors When Tested

Average Training dBA

38 dB   65 dB   75 dB
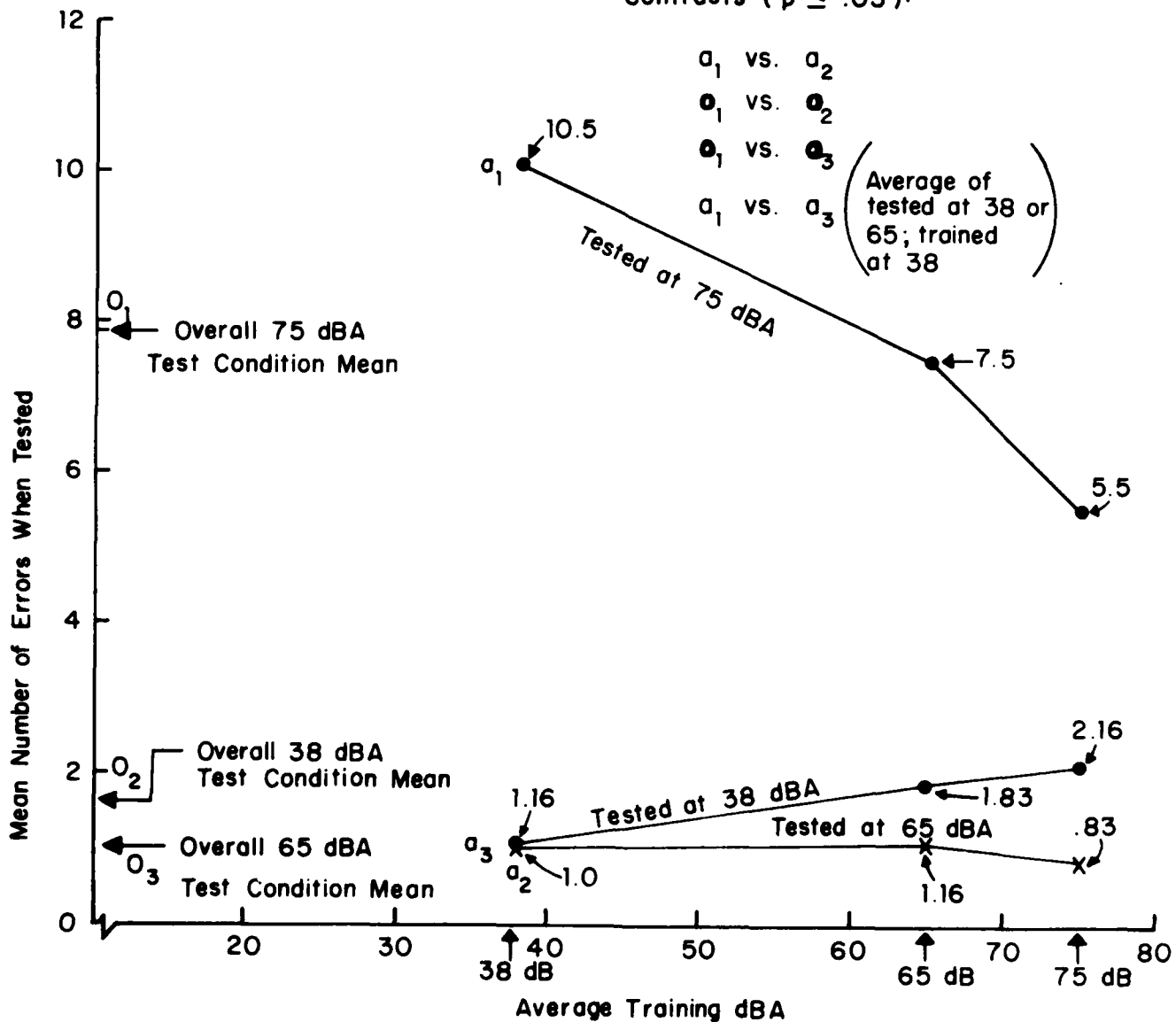
Figure 3.

MEAN NUMBER OF RECOGNITION DEVICE ERRORS
BY TEST AND TRAINING NOISE LEVELS IN
MICROBARS.

(.0002 MICROBARS = THRESHOLD OF HEARING)

17
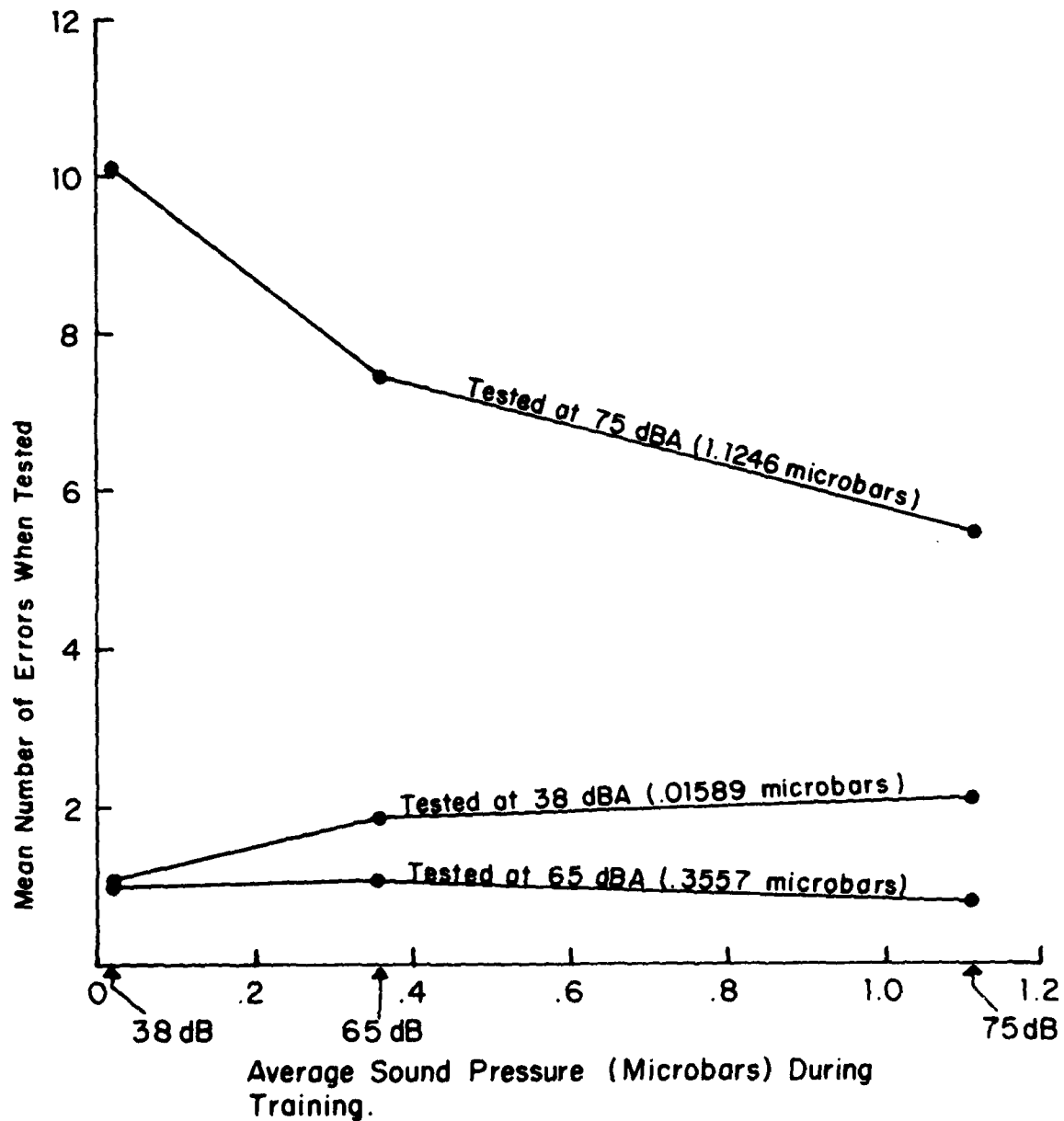
C.   Summary of the Results of the Analysis of Variance.

The results in Table 2 showed that only the F-test for test noise level was statistically significant ($p \ll .01$). Scheffe's contrasts (Table 3) showed that the average number of recognition errors was different ($p < .05$) under the 75dBA testing condition from the average number of recognition errors under either the 38dBA or 65dBA testing condition. These significant differences are shown in Figure 2. Scheffe's contrasts were also used to contrast pairs of test means within and between training conditions. None of the contrasts between pairs of means from different training and different noise level conditions was statistically significant ($\alpha = .05.$). Within training conditions, the only pair of means that was significantly different ($\alpha = .05$) was within the training at 38dBA condition: The means from testing at 75dBA and 65dBA (when trained at 38dBA) were significantly different. Additionally, within the 38dBA training condition, the overall average of the 38dBA and 65dBA average numbers of errors was significantly less than the average number of errors made under the 75dBA testing condition.

Using the nomenclature of Table 1, the joint mean of cells 31 and 32 was not significantly different from the mean of cell 21. In other words, the average of the two high points on the 75dBA testing line in Figure 2 does not differ significantly ($\alpha = .05$) from the low point on the 65dBA line.

14

A second analysis of variance was conducted using a slightly different dependent variable from the one used in the analyses reported in the preceding paragraph. In the second analysis of variance, the type of error labeled extra outputs was excluded from the data, leaving only wrong outputs and "beeps" in the dependent variable. This was done because different microphone utilization practices, or use of a better sound cancelling microphone, might reduce or eliminate extra outputs. The data and the analysis of variance table for this dependent variable excluding extra outputs are given in Appendix IV. Suffice it to say, removing the extra output errors did not change the results of the analysis of variance.

V DISCUSSION AND CONCLUSIONS

The results from the experiment reported here indicate that only the noise condition during testing influenced the number of errors made by the Model T600 Threshold Technology, Inc. voice recognition device. Unlike the results obtained by Drennen, no interaction was found between testing and training background noise levels and number of errors made by the voice recognition device. It should be noted that the sound pressure levels used in this experiment (38, 65, or 75dBA) did not approach the sound intensity levels used by Drennen (10, 101, 106, or 114dB). Drennen (reference note 2) does not consider the results of this experiment to be in conflict with the results he obtained, because he believes the interaction between testing and training background noises

will not be evident until dB levels of around 100 or more are used.

The results of this experiment indicate that care must be exercised if the Model T600 Threshold Technology, Inc. voice recognition device and Shure SM10 microphone are used in an environment with an average conversational background of 75dBA. Overall, (averaged over the three training noise levels) this experiment indicates a higher error in a 75dBA background noise environment than in either the 38 or 65dBA levels. However, a posteriori tests of the mean numbers of errors showed the only significant difference between the 75dBA test condition line and the other two lines in Figure 2 was at the 38dBA training condition. The null hypothesis of no difference in testing performance at 38, 65 or 75dBA cannot be rejected if the device is trained in either a 65, or 75dBA environment. In brief, the results from this experiment indicate that if the Model T600 Threshold Technology, Inc. voice recognition device and Shure SM10 microphone are to be used in a 75dBA conversational background noise environment, then training in a 65 or 75dBA conversational noise environment will yield fewer errors than will training in a 38dBA white noise environment.

There was no significant difference between the average number of errors made in the 65dBA testing condition versus the 38dBA testing condition. The 38 and 65dBA lines in Figures 2 and 4 represent the mean number of errors obtained from the

experiment, but the difference between pairs of 38 and 65dBA
means are not statistically significant - - despite what might
be concluded from casually viewing those lines in Figures 2
and 3.

VI   POSSIBLE FUTURE RESEARCH

Many other possible experiments were suggested during the
course of the experiment described in this report.   The fol-
lowing are suggestions for future experiments.

- The effects of more extreme dB levels of
  background noise on performance of the
  speech recognizer should be determined.

- The effects of background sounds that in-
  clude utterances to which the recognizer
  has been trained should be examined.

- The effects of different kinds of back-
  ground noises, e.g., impact sounds, should
  be studied.

- The effects of different background noise
  levels when different noise cancelling
  microphones are used, and the effects of
  different adjustments to the recognizer
  should be determined.

- The effects of differences among users
  should be studied. (It was noted during
  this experiment that one subject had
  difficulty raising his voice to a level
  comparable to, or above that of, the 75dBA
  background noise.)

- The effects of training of users should
  be ascertained. Can users be trained to
  perform in ways that will maintain system
  performance under different background
  noise conditions?

- Experiments should be conducted in typical
  command and control types of rooms, com-
  partments, etc., as sound reverberations
  in such locations may influence the per-
  formance of a voice recognition system.
  (The experiment described in this report
  was conducted in a soundproof room, which
  also allowed few sound reflections within
  the room.)

- An experiment should be conducted to deter-
  mine if training in a low dBA (e.g., 38dBA)
  conversational environment (if such a low
  dBA conversational environment can be devel-
  oped) has the same effect on performance of
  the recognizer as does training in a low
  intensity white noise environment.

FOOTNOTES

[1] Speech recognition devices are "trained" to recognize selected utterances made by a person. The device is put in a learning mode, and the person repeats the particular utterance a number of times. The device can then be tested to determine if it recognizes the utterance.

[2] The conversational noises were recorded as about twenty people in a room talked informally. They were unaware they were being recorded. For purposes of the experiment, a several minute segment of the original recording was re-recorded to yield a thirty minute length tape. The result of this process was a fairly constant hub-bub of voices, with recognizable words, but no recognizable conversations. The desired level was attained by adjusting gain on an amplifier.

[3] Within the same training noise level, the confidence interval for contrasts between a pair of mean was:

$$C = \sqrt{(IJ\text{-}1) \left\{ F_{.95} \ [(IJ\text{-}1), I(J\text{-}1)(K\text{-}1)] \right\} \ \times \ \frac{2MS}{K} \ [\text{test level} \times \text{Subs (Train. Level)}]}$$

$$C = \sqrt{8 \times 2.27 \times \frac{2}{6} \times 14.48} = 9.35.$$

[4] The confidence interval for contrasting pairs of means from different training and different testing conditions was:

$$C = \sqrt{(IJ\text{-}1) \left\{ F_{.95} \ (1,45) \right\} \ \times 2 \ \left(\frac{(J\text{-}1) \ MS \ (BC \ (A)) \ + \ MS \ (C \ (A))}{JK} \right)}$$

$$C = \sqrt{(9\text{-}1) \times 4.06 \times 2 \left(\frac{2 \times 14.48 + 13.87}{3 \times 6} \right)} = 10.11$$

FOOTNOTES

[4]cont'd

The degrees of freedom for the denominator of the F statistic were computed (Reference note 1.) from:

$$DF_D = \frac{\left\{ \frac{2(J-1)}{JK} \times MS[BC(A)] + \frac{2MS[C(A)]}{JK} \right\}^2}{\frac{\left[\frac{2(J-1)}{JK} [MS[BC(A)]] \right]^2}{I(J-1)(K-1)} + \frac{\left[\frac{2}{JK} \right]^2 \times [MS[C(A)]]^2}{I(K-1)}}$$

$$DF_D = 44.9 \longrightarrow 45$$

The confidence interval for contrasting the combined average of the number of errors in cell 11 (see Table 1) and cell 12 with the average number of errors in cell 13 was:

$$C = \sqrt{(IJ-1) \left\{ F_{.95}[(IJ-1), I(J-1)(K-1)] \right\} \times MS[BC(A)] \left[\frac{(2)^2}{K} + \frac{(-1)^2}{K} + \frac{(-1)^2}{K} \right]}$$

$$C = 16.62$$

24

REFERENCES

Bragdon, Clifford R. Noise Pollution: The Unquiet Crisis. University of Pennsylvania Press, 1971.

Human Engineering Guide to Equipment Design (revised edition), edited by H.P. VanCott and R.G. Kinkade, Washington, D.C.: American Institute for Research, 1972.

Human Factors Engineering Design for Army Materiel, MIL-HDBK-759, 12 March 1975, Commander U.S. Army Missile Command, Standard-ization Division, Redstone Arsenal, Alabama 33809.

Poock, Gary K. Experiments with Voice Input for Command and Control: Using Voice Input to Operate a Distributed Computer Network, Naval Postgraduate School, Monterey, CA, April 1980, Technical Report NPS55-80-016.

Voice Interactive Systems: Applications and Payoffs, Voice Technology for Systems Applications Subgroup of the Depart-ment of Defense Human Factors Engineering Technical Advisory Group, conference held 13-15 May 1980.

Winer, B.J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

Woodworth, R.S. and H. Schlosberg, Experimental Psychology, (Revised edition), New York: Holt, Rinehard & Winston, 1962.

REFERENCE NOTES

1. Collier, Raymond O. Lecture notes from Educational Psychology 218 and 219, University of Minnesota, spring and fall academic quarters, 1964.

2. Drennen, Thomas G. McDonnell Douglas Astronautics Company, St. Louis, Missouri. Phone conversation with author on 12 August 1980.

APPENDIX I

The 50 Utterances Used in the Experiment

Word #	Utterance	Word #	Utterance
0	GRID	25	FIRE
1	LAUNCH	26	TIME
2	COURSE	27	MAP
3	GOLF	28	SCOPE
4	SPEED	29	MAINE
5	MESSAGE	30	NEUTRAL
6	ORDERS	31	REFUEL
7	PLATFORM	32	WHISKEY
8	SENSOR	33	LIMA
9	MISSILE	34	LOGOUT
10	SATELLITE	35	TRACK UNKNOWN
11	NEGATIVE	36	LONGITUDE
12	SUBMARINE	37	TORPEDO
13	ENEMY	38	BLUE FORCE ONE
14	EXECUTE	39	ROMEO
15	SAN FRANCISCO	40	FLIGHT CONTROLLER
16	HUMAN FACTORS	41	SEA OF JAPAN
17	UNITED STATES	42	HONOLULU
18	CLOSE OUT CHARLIE	43	ADVANTAGES
19	COLORADO	44	CONTINUOUS
20	CONNECT TO CHARLIE	45	TASK FORCE COMMANDER
21	NORTH ATLANTIC MAP	46	NORTH CAROLINA
22	COMMAND AND CONTROL	47	BEARING AND DISTANCE
23	CONTINUOUS SPEECH	48	PLOT ALL SUBMARINES
24	VOICE TECHNOLOGY	49	UNITED AIR LINES

turn on machines
load & remove T600 tape
record subjects name, etc. on data collection sheet

APPENDIX II

EXPERIMENTAL PROTOCOL AND SUBJECTS' INSTRUCTIONS

THIS IS AN EXPERIMENT DESIGNED TO EVALUATE SOME VOICE
RECOGNITION EQUIPMENT. I WISH TO EMPHASIZE THAT YOU
ARE NOT BEING EVALUATED - - IT IS THE EQUIPMENT THAT
IS BEING EVALUATED.

THERE ARE TWO DISTINCT PHASES TO THIS EXPERIMENT. IN
THE FIRST PHASE, YOU WILL TRAIN THE EQUIPMENT TO RECOGNIZE
50 UTTERANCES - - AN UTTERANCE BEING A SINGLE WORD OR
SEVERAL WORDS. THE TRAINING MAY BE DONE UNDER A BACKGROUND
NOISE CONDITION. IN THE SECOND PHASE OF THIS EXPERIMENT,
WE WILL TEST THE MACHINE TO SEE IF IT RECOGNIZES YOUR
VOICE. THE TEST WILL BE CONDUCTED UNDER THREE DIFFERENT
BACKGROUND NOISE CONDITIONS. TO SUMMARIZE, WE ARE EVALU-
ATING THE VOICE RECOGNITION EQUIPMENT BY HAVING YOU TRAIN
IT TO RECOGNIZE 50 UTTERANCES. THE TRAINING WILL BE DONE
UNDER ONE BACKGROUND NOISE CONDITION, AND THE TESTING WILL
BE DONE UNDER THREE BACKGROUND NOISE CONDITIONS.

DURING THE TRAINING PHASE, THE UTTERANCES WILL APPEAR ONE
AT A TIME ON THE SCREEN. THE UTTERANCES ARE ALSO ON THIS

**Show subject the
list of utterances.**
PAPER. YOU WILL BE DIRECTED TO REPEAT EACH UTTERANCE 10
TIMES. ATTEMPT TO VARY THE WAY YOU PRONOUNCE AND GIVE
EMPHASIS TO DIFFERENT PARTS OF EACH UTTERANCE. BECAUSE
YOU ARE TO REPEAT EACH UTTERANCE 10 TIMES, YOU MAY FIND
IT USEFUL TO COUNT THE REPITITIONS ON YOUR FINGERS, OR
TO USE CLUSTERS OF, SAY, 3 UTTERANCES TO ALLOW YOU TO
KEEP TRACK OF THE NUMBER OF TIMES YOU HAVE MADE AN UTTERANCE.

TRY TO KEEP THE MICROPHONE IMMEDIATELY IN FRONT OF YOUR

Put headset on subject.
LIPS AND CLOSE TO YOUR LIPS. THERE IS AN ON-OFF SWITCH
FOR THE MICROPHONE. WHEN YOU ARE NOT TRAINING THE MACHINE,
THE SWITCH SHOULD BE OFF. REMEMBER TO VARY THE WAY YOU
PRONOUNCE AND PHRASE THE UTTERANCES.

APPENDIX II

THE 50 UTTERANCES ARE ON THIS LIST. WE'LL SIMPLY TRAIN THEM IN THE ORDER THEY APPEAR ON THE LIST. WE'LL CHECK THEM OFF AS WE GO ALONG.

Turn on noise cassette if appropriate; adjust dB.
WE'LL NOW HAVE YOU TRAIN THE UTTERANCES. (FIRST I'LL TURN ON SOME BACKGROUND NOISE.)

Type: Control U/Return/ Word No.
Check off the words.
Turn off the cassette.
TRAIN UTTERANCES (Test following the training of each word. (Requires two-out-of-three recognition accuracy.)

YOU HAVE NOW FINISHED THE MOST TIME CONSUMING SEGMENT OF THE EXPERIMENT. THE REMAINDER OF THE EXPERIMENT WILL GO RATHER QUICKLY.

WE'LL NOW TEST THE MACHINE'S ABILITY TO RECOGNIZE YOUR UTTERANCES. YOU'LL BE ASKED TO READ OUT-LOUD THE 50 UTTERANCES THREE TIMES - - EACH TIME UNDER A DIFFERENT BACKGROUND NOISE CONDITION.

AFTER THE BACKGROUND NOISE BEGINS, MAKE SURE YOU HAVE THE MICROPHONE SWITCH TURNED ON, AND THEN READ THROUGH THE LIST OF 50 UTTERANCES. PAUSE SEVERAL SECONDS AFTER EACH UTTERANCE, AND I MAY ASK YOU TO PAUSE EVEN LONGER IF I GET BEHIND IN RECORDING ERRORS MADE BY THE EQUIPMENT.

Rewind cassette.
Turn on noise cassette, if appropriate; adjust dB.
Type: Control W/Return
Record errors/turn cassette off.

FIRST TEST

PLEASE READ THE 50 UTTERANCES.

WE'LL NOW REPEAT THE PROCEDURE UNDER A DIFFERENT BACKGROUND NOISE CONDITION.

Rewind cassette.
Turn on noise cassette, if appropriate; adjust dB.
Record errors
Turn cassette off.

SECOND   TEST

PLEASE READ THE 50 UTTERANCES.

WE'RE NOW READY FOR THE LAST TEST.

Rewind cassette.
Turn on noise cassette, if appropriate; adjust dB.
Record errors/ turn cassette off
and rewind

THIRD TEST

PLEASE READ THE 50 UTTERANCES.

28

Recognition Errors by Training and Testing Conditions, and by Subject

Training Noise Level

Testing Noise Level	Subject	1	2	3 (38dB)	4	5	6	7	8	9 (38dB)	10	11	12	13	14	15 (38dB)	16	17	18
38dB		ØE ØB	ØE ØB	2E ØB	3E 1B	ØE ØB	1B ØE	2E ØB	ØE ØB	2E ØB	2E 2B	2E ØB	1B ØE	1E ØB	5E ØB	1E ØB	2E 2B	ØE ØB	1E 1B
65dB		1B 1E	1E ØB	2E ØB	ØE ØB	1E ØB	ØE ØB	1E ØB	ØE ØB	1E ØB	1E ØB	3E ØB	3B ØE	ØE ØB	3E ØB	1E ØB	ØE ØB	ØB ØE	1B ØE
75dB		6B ØE	17E 8B	4E ØB	8E 4B 3+	4E 3B	2E 4B	7E 1B	2E 1B	3E ØB	6E ØB 1+	10E 2B 4+	8B ØE	ØE 1B	3E ØB	3E 1B	2E 1B	11E ØB 5+	4E ØB 2+

E = Wrong Output
B = Beep
+ = Extra Output

29

APPENDIX IV

Errors (wrong outputs & "beeps") by Training Level, Testing Level, and by Subject

Training Noise Level(dBA)

			38						65						75				Overall Mean For a Testing Noise Level
Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Testing Noise Level (dBA) 38	0	0	2	4	0	1	2	0	2	4	2	1	1	5	1	4	0	2	1.72
65	2	1	2	0	1	0	1	0	1	2	3	0	0	3	1	0	0	1	1.00
75	6	25	4	12	7	6	8	3	3	6	12	8	1	3	4	3	11	4	7.00
Overall Mean For a Training Noise Level			4.05						3.22						2.44				3.24 Grand Mean

Summary of the Analysis of Variance

Source of Variance	dF	MS	F
Between Subjects	17		
Training Noise Level	2	11.69	1.17
Subjects within Training Level	15	9.96	-
Within Subjects	36		
Test Noise Level	2	193.05	17.37*
Test Level x Training Noise Level	4	19.39	1.74
Test Level x Subjects within Training Level	30	11.11	-

* $p \ll .01$.

30

INITIAL DISTRIBUTION LIST

No. Copies

1. Defense Technical Information Center   2
   Cameron Station
   Alexandria, VA 22314

2. Library, Code 0142   2
   Naval Postgraduate School
   Monterey, CA 93940

3. Library, Code 55   1
   Naval Postgraduate School
   Monterey, CA 93940

4. Office of Research Admin.   1
   Code 012A
   Naval Postgraduate School
   Monterey, CA 93940

5. Naval Postgraduate School   50
   Monterey, CA 93940
   ATTN: J. Arima, Code 54AA
          R. Elster, Code 54EA
          W. Moroney, Code 55MP
          D. Neil, Code 55NI
          G. Poock, Code 55PK

6. Thomas Brendgord   1
   American Sterilizer Company
   2424 West 23rd St.
   Erie, PA 16152

7. Donald L. Parks   1
   Crew Systems Technology
   Boeing Commercial Airplane Co.
   P.O. Box 3707
   MS 47-08
   Seattle, WA 98124

8. DIPL. Ing. Hartmut Mutschler   1
   Wissenschftl. Mitarbeiter
   Fraunhofer-Institute Fur
   Informations-Und Datenverabeitung
   Rintheimer Strasse 19
   D-7500 Karlsruhe 1
   West Germany

9. J. Mike Byrd   1
    General Dynamics
    Mail Zone 8227-I
    P.O. Box 85106
    San Diego, CA   92138

10. Robert L. Starky   1
    Nuclear Power Generation Division
    Babcox and Wilcox
    P.O. Box 1260
    Lynchburg, VA   24505

11. Barry Drake   1
    Human Factors
    ITT
    Great Eastern House
    Edinburgh Way
    Harlow, Essex
    England

12. Kenneth C. Bice   1
    Texas Instruments
    P.O. Box 2909
    MS 2201
    Austin, TX   78769

13. Walt Goede   1
    Consultant
    31051 Hawksmoor Drive
    Rancho Palos Verdes, CA   90274

14. H. Rudy Ramsey   1
    ITT
    1000 Oronoque Lane
    Stratford, CT   06497

15. Ivan Belyea   1
    Lear Siegler, Inc.
    4141 Eastern Avenue S.E.
    MS 128
    Grand Rapids, MI   49508

16. C.E. (Ned) Wilkins   1
    MTS-MAN-MACHINE INTERFACE
    System Design Dept.
    TRW
    One Space Park
    Redondo Beach, CA   90278

17. Matthew F. Carroll   1
    TRW
    Systems Analysis Dept.
    MS: 75-1900
    One Space Park
    Redondo Beach, CA 90278

18. R. H. Cochrane   1
    Eng. Mgr-Human Factors
    AT&T Long Lines
    Room 4C154
    Bedminster, N.J. 07921

19. Roger L. Kuhn   1
    Director-Ergonomics
    Health Services
    Industrial Relations Staff
    General Motors Corp.
    3044 West Grand Blvd.
    Detroit, MI 48202

20. Jeffrey E. Miller   1
    Northrop Corp.
    Electronics Div.
    2301 West 120th St.
    Hawthorne, CA 90250

21. Michael L. Schneider   1
    Sperry UNIVAC
    Computer Scientist
    P.O. Box 500
    Blue Bell, PA 19424

22. Richard N. Armstrong   1
    U.S. Army Human Eng. Lab
    Box 476
    Fort Rucker, AL 36362

23. Gayle L. Berry   1
    Industrial and Systems Eng
    Ohio State University
    1971 Neil Avenue
    Columbus, OH 43210

24. David J. Cochran   1
    Industrial Engineering
    University of Nebraska-Lincoln
    Lincoln, NE 68588

25. Chris Hale   1
    SRL-Human Factors Eng
    2800 Indian Ripple Road
    Dayton, OH 45440

26. Aaron Marcus   1
    Comp Sci and Appl Math Dept
    Lawrence Berkely Lab
    University of California
    Bldg 50B   Room 2238
    Berkeley, CA 94720

27. Larry C. Lamb   1
    Harris Corp
    MS 22/2419
    P.O. Box 37
    Melbourne, FL 32901

28. Rodney Elden   1
    Mgmt Consultant
    5 Middle Road
    Bronxville, N.Y. 10708

29. Ing. L. Van Breda   1
    Defence and Civil Inst.
     of Environ. Medicine
    P.O. Box 2000
    1133 Sheppard Avenue West
    Downsview, Ontario
    Canada M3M 3B9

30. Carl Rosengrant   1
    Code 8141
    NOSC
    San Diego, CA 92152

31. Nancy Woo   1
    R 84-17
    Merck and Co.
    Box 2000
    Rahway. NJ   07065

32. E. L. Wiener   1
    239-3
    NASA
    Moffett Field, CA 94035

33. Anthony Bessacini   1
    Code 3522
    NUSC
    Newport, RI 02840

34. Bruno Beek   1
    RADC/ IRAA
    Griffiss AFB
    Rome, NY 13441

35. John F. Boehm
    DIRNSA   ATTN: R542
    Fort Meade, MD   20755   1

36. R. Breaux
    Code N-711
    NAVTRAEQUIPCEN
    Orlanda, FL   32813   1

37. CDR Paul Chatelier
    OUSD R&E
    Room 3D129
    Pentagon
    Washington, D.C.   20301   1

38. Ralph Cleveland
    NFMSO
    Code 9333
    Mechanicsburg, PA   17055   1

39. Clay Coler
    Mail Stop 23902
    NASA
    Moffett Field, CA   94035   1

40. Dianne Davis
    NUSC
    Code 3522
    Newport, RI   02840   1

41. Edward DeGregario
    NUSC
    Code 3522
    Newport, RI   02840   1

42. Tice DeYoung
    U.S. Army Engineer Topgraphic
    Lab Research Institute
    Fort Belvoir, CA   22060   1

43. Joe Dickinson
    U.S. Army Applied Tech Lab
    Fort Eustis, VA   23662   1

44. Harold C. Glass
    U.S. Postal Lab
    11711 Parklawn Drive
    Rockville, MD   29852   1

45. Henry Halff
    Code 458
    Office of Naval Research
    Arlington, VA 22217   1

46. Lt. Steve Harris
    6C21
    NADC   1
    Warminster, PA 18974

47. Warren G. Lewis
    NOSC
    Code 8231   1
    San Diego, CA 92152

48. John T. Masterson
    U.S. Postal Lab
    Rockville, MD 20852   1

49. Don McKechnie
    AFAMRL
    Wright Patterson AFB, OH 45433   1

50. Thomas J. Moore
    AFAMRL/BBA
    Wright Patterson AFB, OH 45433   1

51. CAPT. Vince Mortimer
    AFAMRL/BBM
    Wright Patterson AFB, OH 45433   1

52. James Mosko
    Acoustical Sciences Div.
    NAMRL   1
    NAS
    Pensacola, FL 32504

53. CAPT. Leslie K. Scofield
    Directorate of Training
    U.S. Army Signal Center   1
    Fort Gordon, GA 30905

54. C. Skriver
    6021
    NADC   1
    Warminster, PA 18974

55. S. Nils Straatveit
    NUSC
    Code 317   1
    New London, CT 06320

56. Leahmond Tyre   1
    FLTMATSUPPOFFICE
    Code 9333
    Mechanicsburg, PA   17055

57. Eric Werkowitz   1
    AFFDL/FGR
    Wright Patterson AFB, OH   45433

58. T. Weiner   1
    Code 4043
    NADFC
    Warminster, PA   18974

59. Lt. Jeff Woodard   1
    FADC/IRAA
    Griffiss AFB
    Rome, NY   13441

60. Irv Alderman   1
    ARI
    5001 Eisenhower Avenue
    Alexandria, VA   22333

61. John Phillips   1
    Code 7232
    NOSC
    San Diego, CA   92152

62. Wayne Lea   1
    889 Sanford Court
    Santa Barbara, CA   93111

63. Frank Deckelman   1
    Code 330
    NAVELEX
    2511 Jefferson Davis Highway
    Arlington, VA   20360

64. Bruno Witte   1
    Code 8302
    NOSC
    San Diego, CA   92152

65. William Dejka   1
    Code 8302
    NOSC
    San Diego, CA   92152

66. LCDR J. Dietzler   1
    ARPA-IPTO
    1400 Wilson Blvd.
    Arlington, VA   22209

67. Richard Pew
    BBN
    50 Moulton Street
    Cambridge, MA   20138   1

68. Tom Imperato
    5525 Wilshire Blvd.
    Suite 205
    Los Angeles, CA   90036   1

69. Marlin Thomas
    I.E. Dept.
    113 Electrical Eng. Bldg.
    University Missouri-Columbia
    Columbia, MO   64211   1

70. M. Tolcott
    Office of Naval Research
    Code 455
    800 North Quincy Street
    Arlington, VA   22217   1

71. G. Malecki
    Office of Naval Research
    Code 455
    800 North Quincy Street
    Arlington, CA   22217   1

72. N. Greenfield
    BBN
    50 Moulton Street
    Cambridge, MA   20138   1

73. R. Bisbey
    USC-ISI
    Information Sciences Inst.
    4676 Admiralty Way
    Marina Del Rey, CA   90291   1

74. R. Kolb
    NOSC
    Code 824
    San Diego, CA   92152   1

75. CDR R. Meinhold
    CINCPACFLT
    Box 6
    Pearl Harbor, HI   96860   1

76. Daniel Sagalowicz
    SRI
    AI Center
    33 Ravenswood Avenue
    Menlo Park, CA   94025   1

77. Russ Hammond   1
     SAI
     1911 N. Fort Myer Drive
     Suite 1200
     Arlington, VA 22209

78. Stu Parsons   1
     19740 Via Escuela Drive
     Saratoga, CA 95070

79. Don Chaffin   1
     Industrial Eng. Dept.
     University of Michigan
     Ann Arbor, MI 48104

80. Walt Hancock   1
     Industrial Eng. Dept.
     University of Michigan
     Ann Arbor, MI 48104

81. Dennis McCall   1
     NOSC
     Code 8242
     San Diego, CA 92152

82. Marvin Denicoff   1
     Office of Naval Research
     Code 437
     800 N. Quincy Street
     Arlington, VA 22217

83. John Schill   1
     NOSC
     Code 8123
     San Diego, CA 92152

84. H. Morgan   1
     Wharton School
     University of Pennsylvania
     Room W-83
     Dietrich Hall
     Philadelphia, PA 19104

85. Dan Schutzer   1
     NAVELEX
     PME 108
     2511 Jefferson Davis Highway
     Washington, D.C. 20360

86. R. Kahn   1
     ARPA-IPTO
     1400 Wilson Blvd.
     Arlington, VA 22209

87. Tom Martin
     TTI
     18?? Underwood Blvd.
     Pelham, NJ 08075   1

88. Craig Fields
     ARPA-CTO
     1400 Wilson Blvd.
     Arlington, VA 22209   1

89. A.L. Slafkosky   1
     Scientific Advisor
     Commandant of the Marines
     Code RD-1
     Washington, D.C. 20380

     ?????? ???????   1
     Inst. for Defense Analysis
     800 Army-Navy Drive
     Arlington, VA 22202

     Glen Allgaier   1
     Code ????
     NOSC
     San Diego, CA 92152

     Don ??miller   1
     AFAMRL/HEF
     Wright Patterson AFB, OH 45433

     R. Ersatila   1
     Code ??
     NAVELEX
     ???? Jefferson Davis Highway
     Arlington, VA 20360

     ?. ??????   1
     ???? ???
     ?????? Davis Highway
     Arlington, VA 20360

     ????? ?. ??mstrong   1
     ???? ???????? Way
     ???????, Ontario
     Canada ??? ??1

     ?????? ?. Wellington   1
     ???? ??? ???. ?
     Dept. of National Defense
     ??? ?????? Dy Drive
     Ottawa, Ontario
     Canada K?A K?

     R. ??????   1
     RAD?/????
     Griffiss AFB
     Rome, ?? ??? ???

                         46

DATE
ILME